QoS-aware MAC protocols for wireless sensor networks: A survey

M. Aykut Yigitel *, Ozlem Durmaz Incel, Cem Ersoy

Article history:
Received 10 August 2010
Received in revised form 7 February 2011
Accepted 10 February 2011
Available online 16 February 2011
Responsible editor: M.C. Vuran

Keywords:
QoS
QoS challenges
QoS perspectives
QoS mechanisms
Priority assignment
Service differentiation mechanisms
Wireless sensor networks
MAC layer
QoS-aware MAC protocols
Survey

Abstract
The adoption of wireless sensor networks by applications that require complex operations, ranging from health care to industrial monitoring, has brought forward a new challenge of fulfilling the quality of service (QoS) requirements of these applications. However, providing QoS support is a challenging issue due to highly resource constrained nature of sensor nodes, unreliable wireless links and harsh operation environments. In this paper, we focus on the QoS support at the MAC layer which forms the basis of communication stack and has the ability to tune key QoS-specific parameters, such as duty cycle of the sensor devices. We explore QoS challenges and perspectives for wireless sensor networks, survey the QoS mechanisms and classify the state of the art QoS-aware MAC protocols together with discussing their advantages and disadvantages. According to this survey, we observe that instead of providing deterministic QoS guarantees, majority of the protocols follow a service differentiation approach by classifying the data packets according to their type (or classes) and packets from different classes are treated according to their requirements by tuning the associated network parameters at the MAC layer. Design tradeoffs and open research issues are also investigated to point out the further possible research directions in the field of QoS provisioning in wireless sensor networks at the MAC layer.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Wireless sensor networks (WSNs) have appeared as one of the emerging technologies that combine automated sensing, embedded computing and wireless networking into tiny embedded devices. While the early research on WSNs has mainly focused on monitoring applications, such as agriculture [1] and environmental monitoring [2], based on low-rate data collection, current WSN applications can support more complex operations ranging from health care [3] to industrial monitoring and automation [4]. Besides these, the availability of low-cost hardware and rapid development of tiny cameras and microphones have enabled a new class of WSNs: multimedia or visual wireless sensor networks [5,6] and this new class has contributed to new potential WSN applications, such as surveillance.

What is common in these emerging application domains is that performance and quality of service (QoS) assurances are becoming crucial as opposed to the best-effort performance in traditional monitoring applications.

The term QoS is widely used in the area of all kinds of networks but still there is no consensus on its exact meaning. International Telecommunication Union (ITU) Recommendation E.800 (09/08) has defined QoS as: “Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service”. Traditionally it refers to the control mechanisms that orchestrate the resource reservation rather than the provided service quality itself. Simply or practically, QoS brings the ability of giving different priorities to various users, applications, and data flows, frames or packets based on their requirements by controlling the resource sharing. Hence higher level of performance over others can be provided through a set of measurable service parameters such as delay, jitter, available bandwidth, and packet loss.
QoS requirements in traditional data networks fundamentally stem from the end-to-end bandwidth-hungry multimedia applications [7]. In this context, reservation-based approaches, such as Integrated Services or IntServ [8], are widely used in providing QoS guarantees. However, guaranteeing a certain QoS is a challenging issue due to the unpredictable nature of the wireless links, unstable topology (due to node failure or link failure) and severe resource constraints in WSNs. These constraints make it harder to adopt the existing solutions in wired and other wireless networks. Besides these constraints, while recent applications, especially real-time, multimedia and mission-critical applications, call for QoS support, the inherent characteristic of WSNs, “energy efficiency” makes the QoS provision a challenging task.

Parallel to recent advancements, WSN applications have become more and more bandwidth-hungry and delay-sensitive. In order to meet these requirements, WSNs need novel and well-designed QoS support in each layer of the communication protocol stack since envisioned applications are dissimilar to traditional end-to-end applications. Especially real-time multimedia and mission-critical applications brought forward new QoS requirements since they need delay-bounded and reliable data delivery. This variety of the applications and requirements of these applications make implementation of a “one-size-for-all” QoS-support mechanism impossible. However, well-defined requirements and QoS parameters can be a guide to develop QoS-support for effective and efficient delivery of sensor data.

In this work, we focus on the QoS support at the MAC layer and survey the existing protocols in the literature. Although centralized MAC schemes exist for other types of networks, such as point coordination function (PCF) in IEEE 802.11, where nodes request the right for medium access from a coordinator, these schemes are hardly applied to WSNs due to the large number of sensor nodes, multi-hop nature of the networks and scalability issues. Therefore, our focus is on distributed QoS support at the MAC layer. The reason why we focus on the MAC layer is that, all other upper-layer components are dependent on the MAC layer and this makes it a primary decisive factor for the overall performance of the network. Nowadays, cross-layer solutions for WSNs where functionalities of multiple traditional layers are melted into a functional module, are widely adopted [9]. By the cross-layer approach, a single module can obtain every necessary information regardless of the layer abstraction and has chance to optimize the overall performance of the sensor network. However, interoperability or interchangeability between layers cannot be mentioned in this case since there is no layer abstraction within the protocol stack. In case of QoS support, there is no distinction between layered and cross-layer protocols. QoS awareness can be adopted with the same goals and challenges by both concepts.

In this paper, our aim is to survey the existing QoS-aware MAC protocols for WSNs including mobile, underground and underwater sensor networks. To the best of our knowledge, although there exist surveys on QoS support in WSNs [7] and on MAC protocols for WSNs [10,11], there is no extensive survey paper on the QoS-aware MAC protocols, including their comparative evaluation. Although, Zogovic et al. [12] briefly summarize QoS Provisioning at MAC and physical layers for WSNs, they neither provide an extensive survey, nor discuss the comparisons and provide a classification together with future research directions.

Our contribution is to present a detailed survey on the topic and discuss the open issues in this domain which, we believe, is going to receive a lot of attention in the coming years. We start with a background information in the context of QoS provision in wired and wireless networks. We summarize different types of QoS approaches and discuss which can be applied to WSNs. Additionally, we mention the QoS perspectives, namely application-specific QoS and network-specific QoS, and discuss the requirements of different types of applications. Then, we elaborate on the challenges of QoS provisioning in WSNs and discuss the QoS metrics, such as bounded delay, guaranteed throughput, together with the tunable parameters at the MAC layer, such as duty cycle, contention window size. After explaining the metrics and parameters, we discuss the QoS mechanisms that can be applied in the context of WSNs. We then continue explaining the details of existing QoS-aware MAC protocols for WSNs including their QoS metrics, parameters, mechanisms and present an extensive comparison of them. We conclude the paper with open research issues and possible future research directions.

The rest of the paper is organized as follows: in Section 2, we provide background information on QoS support in wired and wireless networks. In Section 3, we discuss the QoS challenges and continue with the QoS metrics in WSNs in Section 4. We present the QoS mechanisms in Section 5 and explain the details of the existing QoS-aware MAC protocols in WSNs and give comparisons in Section 6. Section 7 discusses the MAC layer tradeoffs and Section 8 elaborates on the properties of a well-defined MAC protocol. In Section 9 we discuss the open issues and give possible directions for the future research. Finally, in Section 10, we draw the conclusions.

2. Background and QoS perspectives

Internet was initially designed for providing the best effort delivery of application data since average performance guarantees were sufficient for initial types of applications [13]. However, with the emergence of applications, such as Internet telephony and video streaming, that require high throughput, bounded delay, bounded delay jitter, and high reliability, best effort delivery has become insufficient to support these applications. Consequently, this has driven and enabled the development of algorithms, protocols and mechanisms that provide QoS support for diverse set of applications. A similar situation is currently observed in WSNs. Traditionally, WSNs have been used for monitoring applications based on low-rate data collection with low periods of operation. Current WSNs are considered to support more complex operations ranging from target tracking [14] to assisted living [15] which require efficient, reliable and timely collection of large amounts of data. Moreover, the recent advances in image sensor
technology, have enabled the use of video sensors and this resulted in a new class of WSNs, called visual or multimedia sensor networks [5,6], that can be used for various potential applications, such as telepresence and surveillance. It is certain that, these networks also have tighter QoS requirements, such as low data delay and maximum reliability, compared to traditional WSNs [6].

2.1. QoS provisioning and service differentiation in traditional networks

Shortly, QoS is the ability of a network to satisfy the certain requirements of the user or application. There are two main types of QoS provision defined in wired and wireless networks: Hard QoS and Soft QoS. The applications that require hard QoS should be provided deterministic QoS guarantees, such as strict bounds on packet delays, bandwidth or packet losses. In soft QoS approach, again the application has tight QoS requirements but the temporal violations on QoS provisioning can be tolerated to a certain extent [13].

Service differentiation is the widely adopted scheme in both wired and wireless networks to provide hard/soft QoS guarantees. There are two service differentiation models proposed for conventional computer networks, Integrated services (IntServ) [8] and differentiated services (DiffServ) [16]. Aim of both the differentiation models are to prioritize flows or packets, map their priorities into service qualities and provide required service quality by sharing limited resources among them.

IntServ model maintains service on a per-flow basis and can be considered as a reservation-based approach. It specifies a fine grained QoS system and follows the hard QoS approach [17]. Flows can be considered as data-centric or host-centric where data-centric consideration can be information generated by motion sensors from a commonly used breach path in border surveillance and host-centric consideration can be the stream of packets between a particular source and destination. However, IntServ model has a number of disadvantages which makes it inappropriate for WSNs. Firstly, it is hard to provide guaranteed service quality due to time varying channel capacity on the wireless medium. Second, maintenance of the per-flow states of the sensor nodes and scalability for dense networks is a real challenge. Third, IntServ model requires a reliable in-band or out-of-band QoS signaling within the sensor network for resource reservation which is very hard to assure in WSNs.

DiffServ model maintains service on a per-packet basis and can be considered as a reservation-less approach. Major drawback of DiffServ model is its costly memory requirement since every network entity will behave as a source and an intermediate hop. However, lightweight and easy-to-implement DiffServ model can be adapted to WSNs easily and this model operates in a multi-hop manner [18]. Each packet will have a degree of importance and this will be apparent for every entity of the network. In this way, each layer of the communication protocol stack can treat the packet by the way its priority imposes. Therefore, DiffServ model will be assumed as the default service differentiation method for the rest of our work.
control. Therefore, the MAC layer plays a key role for QoS provisioning and dominates the performance of the QoS support. The reader can refer to [24–29] for QoS support at the network layer, and to [30–33] at the transport layer and to [34] for different layers.

3. QoS challenges in WSNs

WSNs inherit most of the well-known QoS challenges from traditional wireless networks, such as time varying channels and unreliable links [35]. However, typical characteristics of WSNs, such as severe resource constraints and harsh environmental conditions, pose additional unique challenges for QoS-support. These QoS challenges for WSNs are explained in this section:

- **Resource constraints**: WSNs lack of bandwidth, memory, energy and processing capability. However, limited energy is the most crucial one since in many scenarios it is impossible or impractical to replace or recharge batteries of the sensor nodes. Although energy harvesting via solar energy [36,37] seems to be a promising solution to energy scarcity, present solar panels are still too large for tiny sensor devices. Eventually, proposed QoS support mechanisms must be lightweight and simple in order to operate on a highly resource constrained sensor node.

- **Node deployment**: Deployment of the sensor nodes may be either deterministic or random. In deterministic deployment, sensor nodes are placed by hand and routing can be performed through pre-scheduled paths. In a random deployment, sensor nodes are deployed randomly and organize themselves in an ad hoc manner. Hence, neighbor discovery, path discovery, geographical information of the nodes and clustering are the issues to be solved.

- **Topology changes**: Node mobility, link failures, node malfunctioning, energy depletion or natural events like flood or fire can cause topology changes. Moreover, most of the link layer or MAC layer protocols employ sleep-listen schedules and turn the radio of the sensor nodes off temporarily for energy saving. This kind of power management mechanisms also cause frequent topology changes. Inevitably, dynamic nature of the WSN topology introduces an extra challenge for QoS support.

- **Data redundancy**: WSNs comprise a large amount of tiny sensor nodes and hence, observed event or phenomena can be detected by several sensor nodes. Although this redundancy helps reliable data transfer, it also causes unnecessary data delivery in the network which consequently yields to congestion. Data aggregation/fusion [38,39] mechanisms may decrease the redundancy but also may introduce additional delay and complexity in the system. Therefore, effective QoS mechanisms are needed to cope with the data redundancy.

- **Multiple traffic types**: Sensor nodes which have the capability of sensing or observing various phenomena can generate different types of traffic. For instance, streaming multimedia and location of a detected target or periodic temperature information of an area might be carried at the same time for a specific application. Therefore, applications requiring existence of multiple traffic classes add extra challenging issues to QoS support since requirements of traffic classes differ from each other.

- **Real-time traffic**: In some critical applications like natural disaster monitoring or security surveillance, gathered data is valid only for a limited time frame and has to be delivered before its deadline. This type of critical real-time data must be handled by adequate QoS mechanisms.

- **Unbalanced traffic**: In a WSN, there is usually a central entity (sometimes multiple of them) that obtains the global view of the sensing environment called the sink node and there may exist middle layer entities for data aggregation and compression named as cluster heads. Therefore, unbalanced traffic flows from sensor nodes to sink nodes or cluster heads are commonly observed in WSNs. Moreover, event-driven applications mostly cause sporadic changes in the traffic pattern in case of event detection. Although smart routing protocols may share the traffic load between different routes, MAC protocol still has to accommodate unbalanced and bursty traffic.

- **Scalability**: Most of the WSNs are composed of hundreds or thousands of sensor nodes. As the area of interest or requirements for the quality of observation increase, more sensor nodes need to be deployed. Therefore, designed QoS mechanism must scale well with highly dense or large scale networks.

Together with successful deployment examples of traditional terrestrial sensor networks, researchers started to work on using sensor networks in different environments such as underwater and underground. Both Underwater Acoustic Sensor Networks (UW-ASNs) [40] and Wireless Underground Sensor Networks (WUSNs) [41] differ from traditional terrestrial sensor networks since they operate in diverse environments and communicate through totally different mediums. The diversities in the operating environment and the communication medium have significant effects on the network itself and therefore, pose some additional challenges for QoS support. Except the ones inherited from traditional terrestrial sensor networks, those additional QoS challenges for UW-ASNs and WUSNs can be listed as follows:

- **Underwater/underground channel**: Both underwater [42] and underground [43,44] channels show significant spatial and temporal differences. Also, the propagation delays in underwater and underground are five orders of magnitude higher than the traditional terrestrial channels. Hence, designed QoS mechanisms must take the highly dynamic nature of the channel into account.

- **Higher error rates**: High bit error rates (BER) can be experienced due to high communication medium density for both water [45] and soil. Moreover, connectivity losses occur more frequently due to heavy multipath and fading. Therefore, effective error control mechanisms must be integrated to achieve acceptable level of BER.
• Extreme environmental conditions: Extreme characteristics of both underwater and underground environments make sensor devices more prone to corrosion and malfunctioning which shortens the network lifetime and decreases the level of reliability. In order to cope with these problems, derived QoS mechanisms must take extreme environmental conditions into account and take the necessary measures beforehand.

In WSNs, sensor nodes are generally assumed to be static. However, some recent applications of WSNs, such as medical care and disaster response, utilize mobile sensor nodes and mobility poses another set of unique challenges to be addressed which include topology management, routing, energy management. Since the neighborhood of a node changes frequently due to the mobility, the topology and spatial density of the network also change frequently. Hence, QoS provisioning in mobile sensor networks becomes a more challenging task since envisioned methods must handle highly dynamic node connectivity and density.

WSN related challenging issues are highlighted. These challenges make it difficult for providing deterministic QoS guarantees, such as strict bounds on packet delays, guaranteed bandwidth or packet losses in WSNs. However, providing different services for different traffic classes in spite of these challenges are still feasible as we further discuss in the rest of the paper. These mentioned challenging factors must be taken into account during the design of new QoS-support mechanisms and novel techniques have to be adopted in order to cope with them.

4. QoS requirements, metrics and parameters

In this section, we first highlight the QoS requirements in WSNs from the perspective of the requirements of different data collection models [46]. Next, we focus on the metrics and parameters to be tuned for QoS provisioning.

4.1. QoS requirements

Although our focus is on network-specific QoS in WSNs, as we mentioned in Section 2.2, QoS requirements of different applications differ from each other. For instance, traditional low-rate data collection applications may tolerate delay and jitter but packet losses may be important for the application whereas high rate, real time applications, such as target tracking, require a bound on the maximum acceptable delay. Therefore, application requirements are also important for network-specific QoS. Rather than investigating the QoS requirements of every application in WSNs, it is a better approach to focus on the data delivery models that are used in different applications and map the requirements of these data collection models to a set of QoS metrics. This approach was also followed in [7]. Depending on the application requirements, there are three basic data delivery models: continuous, query-driven, and event-driven model [46]. In the following part, we discuss these models and their associated QoS requirements:

1. Event-driven: In this model, sensor nodes report data only if an event of interest occurs. Usually, the events are rare. Yet, when an event occurs, a burst of packets are often generated that need to be transported reliably, and usually in real-time, to a base station. The success of the network depends on the efficient detection and notification of the event that is of interest to the user. This is bound to quality and accuracy of the observation related to the observed phenomena with reliable and fast delivery of the information about the detected event. Since more than one sensor nodes will detect the event and generate related data, this type of applications are not end-to-end. Also creation of highly redundant and bursty traffic by sensors affected by the same event is very likely to be observed in event-driven applications. Surveillance and target tracking can be an example for this class.

2. Query-driven: Query-driven data delivery model is very similar to the event-driven model with an exception: Data is pushed to the sink without any demand by the sensor nodes in event-driven model while data is requested by the sink and pushed by the sensor nodes in the query-driven model. Hence, contrary to the one-way traffic of event-driven model, two-way traffic comes into scene which consists of requests of the sink and replies of the sensor nodes. Both requests and replies must be delivered quickly and reliably for achieving higher performance in query-driven applications. Environmental control or habitat monitoring can be an example for this class.

3. Continuous: In this model, sensor nodes transmit the collected data at periodic intervals and can be considered as the basic model for traditional monitoring applications based on data collection. The data rates can be usually low and to save energy the radios can be turned on only during data transmissions if scalar data is collected. However, real-time data such as voice or image are delay-intolerant and requires a certain level of bandwidth. Also packet losses are tolerated in a limited threshold. For periodically collected non-real-time data, latency and packet losses are tolerable. Surveillance or reconnaissance can be an example of this class.

4. Hybrid: If the mentioned data delivery models coexist in the same network, carried traffic must be classified and requirements of these traffic classes must be satisfied. A surveillance application that sends both periodic temperature and event-triggered video data is an example of the hybrid model.

4.2. QoS metrics and parameters

In the previous subsection, we discussed the QoS requirements of WSNs from the perspective of applications that adopt similar data collection models. In this section, we present the metrics that quantify these QoS requirements. The general metrics from the networking perspective are maximizing throughput and goodput, minimizing delay, maximizing reliability, minimizing delay jitter, maximizing energy efficiency, etc. In order to perform well regarding these metrics, the overall impact of the whole
protocol stack should be taken into account while supporting QoS. However, since our focus is on the MAC layer, we focus on the performance metrics that can be fulfilled at the MAC layer, as follows:

- **Minimizing medium access delay:** It is certain that in order to minimize the end-to-end delay from sensor sources to the sink node, the performance of routing layer should also be taken into account. What can be done at the MAC layer in terms of delay is to minimize the medium access delay of the sensor devices to ensure that the packet latency is optimized to meet the end-to-end delay requirements.

- **Minimizing collisions:** Collisions, and consequently retransmissions, directly impact the overall networking metrics such as throughput, delay and energy efficiency. Since the MAC layer coordinates the sharing of the wireless medium, it is responsible for minimizing the number of collisions. Collisions can be prevented by careful carrier sensing methods, such as adapting contention window according to the traffic requirements, considering the contention-based protocols. Similarly, adapting the number of time slots, frequencies according to network requirements can prevent collisions in the case of contention-free protocols.

- **Maximizing reliability:** Related with minimizing the collisions, MAC layer can also contribute to reliability assurance. Acknowledgement mechanisms can be used to identify the packet losses and accordingly retransmissions can be performed in time to fix the problems.

- **Minimizing energy consumption:** Energy efficiency is still the most important requirement in WSNs due to the battery-limited operation of sensor devices. MAC layer can contribute to energy efficiency by minimizing collisions and retransmissions and more importantly can tune the duty cycle of the sensor devices according to the network dynamics. Duty cycling is important in WSN operations since the wireless operation consumes most of the energy and radio should be kept off whenever it is not needed. Moreover, transmission power of the sensor radios can be adapted according to network conditions to minimize energy consumption at the MAC layer.

- **Minimizing interference and maximizing concurrency (parallel transmissions):** Since wireless medium is a shared medium, all unwanted transmissions within the same network or transmissions from other networks that share the same parts of the spectrum contribute to interference on the intended transmissions. Interference causes packet loses and hence affect the throughput, delay and energy efficiency of the network. Maximizing concurrency while limiting the impact of interference on parallel transmissions can contribute to these metrics. MAC layer can achieve minimal interference and maximum concurrency by tuning the related parameters, such as contention window, timing, transmission power, operating channel.

- **Maximizing adaptivity to changes:** WSNs are characterized by their dynamic behavior: nodes may deplete their battery and disconnect from the network, new nodes may be added to the network, links between nodes may change in time due to environmental conditions or topological changes, traffic conditions may change according to the monitored phenomena. Therefore, MAC protocols should take adaptive actions according to the network dynamics. For instance, if high-rate, real-time data traffic dominates in the network nodes should work with a high duty cycle whereas if low-rate traffic flows in the network most of the nodes can be kept as passive to conserve energy.

As we mentioned, these are the metrics that can be fulfilled at the MAC layer whereas other metrics such as maximizing throughput and goodput, minimizing end-to-end delay from sources to the sink node can be considered for the whole protocol stack. In order to fulfill these performance objectives, the associated parameters should be tuned at the MAC layer accordingly. These parameters include transmission power, timing or frequency of transmissions (either with adapting contention window and backoffs in contention-based protocols, or adapting time slots or frequencies in contention-free protocols), duty cycle, queuing mechanisms, acknowledgement mechanisms and bandwidth.

Although MAC related QoS metrics are highlighted, it is not mandatory or practical to provide each of them in a single MAC protocol since requirements of the sensor network applications are utterly different. Therefore, in Table 1, we assign the QoS metrics to the application classes defined in Section 4.1 in order to simplify the requirement-metric matrix. However, both the applications and the metrics are not limited to those listed in this section. Hence, Table 1 does not exhibit an absolute pairing, it is just shows the basic matches.

### 5. QoS mechanisms in WSNs at MAC layer

Although each method contributing to improve the performance of the MAC layer and to fulfill the QoS requirements can be counted as QoS mechanism, there is a bunch of them already proposed and applied in the literature. In this section, properties of these mechanisms and how they provide QoS will be investigated briefly. Examples of QoS-aware MAC protocols in the literature utilizing these techniques will be surveyed in Section 6.

#### 5.1. Adaptation and learning

Adaptation mechanisms at the MAC layer provide QoS by adapting operation parameters of the sensor nodes to

---

**Table 1**

<table>
<thead>
<tr>
<th>Important MAC layer QoS metrics for application classes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>QoS metric</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Medium access delay</td>
</tr>
<tr>
<td>Collision rate</td>
</tr>
<tr>
<td>Reliability</td>
</tr>
<tr>
<td>Energy consumption</td>
</tr>
<tr>
<td>Interference/concurrency</td>
</tr>
<tr>
<td>Adaptivity</td>
</tr>
</tbody>
</table>
the current network conditions according to their local or collaborative observations such as traffic pattern, network topology, collision probability or channel condition. By this way, sensor nodes fine tune their operation parameters such as duty cycle, contention window size, backoff exponent or transmission slot scheduling and try to accommodate offered traffic load in a more efficient way.

Similar to adaptation, sensor nodes may try to learn the characteristics of the network during their operation and take the necessary adaptive precautions against changing conditions beforehand rather than responding afterwards. However, learning algorithms require certain amount of time to make accurate predictions and accuracy of the predictions increases in time. More importantly, envisioned learning algorithms must be simple and lightweight to be used in resource constrained sensor nodes.

5.2. Error control

Aim of the error control mechanisms is to reduce energy consumption while providing reliable and fast delivery of the sensory data. However, error control is not a layer-specific issue and can be implemented in each layer of the communication protocol stack. There are three mechanisms most commonly used for error control: Automatic Repeat Request (ARQ), Forward Error Correction (FEC) and Hybrid ARQ [47].

ARQ scheme can be used to provide guaranteed hard QoS by persistent retransmissions until the data is successfully delivered. However, performance of ARQ is closely related with the channel conditions and probability of collisions. If the channel is in good condition and not overloaded; retransmissions are rarely needed and ARQ can improve successful data delivery ratio significantly. On the contrary; latency, drop ratio and energy consumption per successfully transmitted packet can grow to unacceptable levels, especially for delay-bound real time traffic in case of frequent retransmissions.

The idea behind the FEC mechanism is to prevent retransmission of the entire data packet in case of partial errors by including some redundancy in it. This redundancy is then used to recover failures caused by wireless channel at the receiver side. Redundant data might be additional bits added during source coding or packets added during fragmentation of a video frame. However, the FEC mechanism requires additional memory for data queues and brings an extra latency caused by transmission of longer data packets. Also, the FEC coding algorithm must be lightweight and simple since sensor nodes are equipped with very low clock-rate processors. Although the FEC mechanism has certain shortcomings, they can be alleviated by changing the strength of the FEC code based on the current channel conditions.

Hybrid ARQ takes advantage of both ARQ and FEC mechanisms. Initially, data packets are weakly coded or not coded at all by the sender. If the received packet is in error and cannot be recovered, receiver sends a negative acknowledgement to the sender. The sender than recodes the packet with a more powerful FEC code and resends the packet. This cycle continues until the packet is successfully delivered.

5.3. Data suppression and aggregation

Data suppression and aggregation mechanisms try to minimize radio communication by reducing the traffic load of the network, hence provides energy saving [38]. The redundancy can be eliminated by either suppressing the set of messages belonging to the same event before being transmitted or by combining the data coming from different sources. This elimination also prevents congestions caused by overloading, decreases probability of collision and improves the utilization of the network resources such as bandwidth.

Data suppression and aggregation techniques are strictly application dependent and similar to error control, they can be implemented in any layer of the protocol stack. Although layer arbitration brings modularity and flexibility, cross-layer solutions can improve the Degree of Aggregation (DoA) by exploiting contents of the data semantically. However, there is a tradeoff between energy and latency in data suppression. As the router nodes wait for other packets to aggregate, the latency of the packets being aggregated increases. Meanwhile, this provides extra power conservation by reducing the radio communication. Therefore, the DoA must be retained in a reasonable level without violating the QoS constraints of the data.

5.4. Power control

The main idea of power control is simply adjusting the transmission power of the sensor nodes according to the minimum power required for successful transmission [48]. Many factors affect the required minimal power including frequency of the band, wireless channel conditions (e.g. noise, path loss, shadowing) and distance to receiver. Although power control is a physical layer related issue, it has a significant impact on both MAC and network layers since it has the ability to control the network connectivity. Therefore, the power control mechanism can be implemented in the MAC layer and a joint physical-MAC layer solution can be derived.

We can count the reduction of energy consumption as a primary contribution of power control to QoS provisioning. Also, it increases the concurrent communications by decreasing interference, hence improves the channel utilization. However, dynamic nature of the wireless links makes the implementation of power control mechanism a challenging task.

5.5. Clustering

It is very hard to provide global synchronization in WSNs considering the large deployments and the number of sensor nodes. This challenge has led the development of clustering mechanisms to simplify the synchronization and coordination by grouping set of neighboring sensor nodes. Clustering provides significant energy saving by improving inter-node connectivity and facilitating data aggregation, hence can be used to provide QoS support in terms of energy consumption and reliability. Clustering algorithms can be classified as static and dynamic.
Static clustering algorithms select the head and members of the cluster once during the deployment or initialization phase of the network and the role of the sensor nodes does not change in time. Static prioritization is easy to employ and does not require any control messaging. However, the network lifetime and connectivity can be severely damaged since cluster heads consume more energy and their batteries get depleted earlier.

Dynamic clustering reconstructs the clusters or rotates the cluster heads according to the current topology and tries to distribute the forwarding load evenly among cluster members. Hence, early battery exhaustion of cluster heads can be prevented. However, this method introduces significant overhead due to inter-cluster and intra-cluster control message exchanges.

5.6. Service differentiation

Service differentiation is the most widely known and utilized technique for QoS provisioning not only in WSNs but also in all kinds of wired and wireless networks [18]. However, service differentiation is not the QoS support itself, it is just a mechanism to meet the requirements of the users or applications properly. It differentiates and prioritizes the traffic carried on the network based on one or more criteria and forms several traffic classes. In this way, MAC layer treats each of these traffic classes differently by managing the resource sharing among them and tries to fulfill the requirements imposed by their degree of importance. Thereby, service differentiation consists of two phases: (i) priority assignment; and (ii) differentiation between priority levels.

5.6.1. Priority assignment

Priority assignment methods that imply the criteria of differentiation need to be identified carefully in order to achieve fair and effective QoS support. Since the correctness and accuracy of the assigned priorities affect the QoS support significantly, overall performance of the QoS mechanisms highly depends on it. As mentioned in Section 2.1, reservation-less DiffServ model is in the scope of this paper and priority assignment methods in DiffServ are divided into three categories:

1. Static priority assignment: If the priority is assigned once the packet is created and never changes until its destination, it is called as static priority assignment. Decision parameters for static priority assignment can be listed as follows:
   - Traffic class: Packets can be prioritized based on the type of traffic like real-time, non-real-time, best effort. Accordingly, delay and loss bounded real-time packets will have higher priority whereas non-real-time and best effort packets have lower [49,50].
   - Source type: QoS mechanism can specify set or sets of sensor nodes or sinks which generate more important data than others and assign all network entities a priority. Consequently, the node which generates the packet also gives the priority of itself to its packets, i.e. packet inherits the priority of its creator.

Priorities of the entities can be given based on the sensor type, observed area characteristics, distance to center or sink [51].

- Data delivery model: There are four types of data delivery models in WSNs as discussed in Section 4.1. Priority of the packets can be selected based on the associated data delivery model. For example, event-driven data might have higher priority than periodic messages in case of an intrusion detection application [52].

2. Dynamic priority assignment: Contrary to the static priority assignment, packet priorities may vary during delivery. There are several criteria proposed for dynamic prioritization:
   - Remaining hop count: In a multihop WSN, remaining number of hops to the destination of the packet can be used as a parameter for packet prioritization. One of the ideas behind this parameter is minimizing the delay deviations between the packets generated by the sensor nodes which have different distances to the sink. Also, as the distance that the packet will travel increases, it becomes more vulnerable to deadline miss, dropping and link failure. Hence, packets which will traverse more hops are given higher priority.
   - Traversed hop count: The number of traversed hops can be used for prioritization since losing, dropping or missing the deadline of a packet which has traversed more hops will be waste of more network resources than the one which has traversed less hops. Therefore, giving higher priorities to the more invested packets in terms of network resources increases the network lifetime and channel utilization. Moreover, relatively further sensor nodes from the sink usually have smaller chance to deliver their packets and suffer from high latencies. Hence, speeding up the packet as it gets closer to the sink also provides fairness among sensor nodes in terms of packet delivery ratio and latency. Examples of such dynamic priority assignment schemes can be found in [53,54].
   - Remaining energy: The closer a packet is to miss its deadline, the higher priority it should have, since the packet will be useless after its deadline. In this way, waste of network resources can be prevented [55].
   - Packet deadline: The closer a packet is to miss its deadline, the higher priority it should have, since the packet will be useless after its deadline. In this way, waste of network resources can be prevented [55].
   - Remaining energy: Increasing the priority of the packets as the remaining energy of the generating or relaying sensor node decreases, extends the lifetime of the sensor node by preventing the energy waste caused by idle listening. Examples of protocols that provide differentiation based on remaining energy are presented in [56,54].
   - Traffic load: Forwarding loads of the sensor nodes can change depending on their position or role (leaf node, relay node, cluster head) in the network. Giving higher priority to the sensor nodes that have relatively heavier forwarding load can decrease the packet dropping ratio caused by buffer overflow. Besides its role in the network, proportional buffer
load of the sensor node can be an indicator of the traffic load also [56,54].

3. Hybrid priority assignment: Priority of the packets can be determined in a hybrid manner by considering both static and dynamic decision criteria. Moreover, by giving certain weights to these criteria, importance degree of the packet can be calculated more precisely and mapped to a priority level.

5.6.2. Differentiation methods

After priority assignment, the second and crucial phase of the service differentiation is resource sharing according to the importance of the carried data. There are some techniques at the MAC layer to provide different quality of services to different traffic classes and can be listed as follows:

- **Changing Contention Window (CW) size**: Contention based medium access schemes necessitate a contention period between the sensor nodes that attempt to send data concurrently, in order not to interfere with each other's transmission. Following the contention period, one of these sensor nodes wins the contention and qualifies to reserve the communication channel and sends its data. Since contention period determines the sensor node which will be served next, it has a direct effect on the medium sharing among all sensor nodes. We can extend this medium sharing also among all traffic classes carried in the network if sensor nodes are assessed according to their data waiting to be transmitted. Hence, the desired service quality can be provided to specific traffic classes by favoring the sensor nodes which have data belonging to that particular traffic class during contention period as in [49,51–58,53,56,58]. Traditionally, each contender node sets a timer or selects a contention slot. The first sensor node whose timer expires or whose slot time arrives reserves the medium and starts sending its data. By setting relatively shorter CW sizes for sensor nodes with higher priority traffic, it can be assured that the timer or slot of that sensor node will expire before others. Similarly, setting longer CW sizes for sensor nodes with lower priority traffic decreases their medium reservation chance. This method also has an indirect contribution to more qualified service provisioning by reducing the probability of collision since contentions mostly occur within the reduced set of nodes belonging to the same priority group [59].

- **Changing contention slot selection probability**: In random access MAC schemes, contender nodes normally select a contention slot in a random fashion. However, employing non-uniform probability distributions for contention slot selection makes significant difference [54]. For instance, using a decreasing geometric distribution can increase the chance of medium reservation for a node since smaller contention slots are most likely to be selected.

- **Changing inter-frame space (IFS) duration**: In contention based medium access schemes, IFS is defined as the amount of time that sensor nodes stay quiet just before the contention or backoff period. Employing different IFS values for sensor nodes having different kinds of traffic classes provides service differentiation among them and gives precedence to the ones using shorter IFS [56,53,49,52].

- **Changing backoff exponent**: Although IFS and contention periods are utilized to overcome collisions in contention based medium access schemes, it is impossible to totally eliminate collisions because more than one sensor nodes may set their timers to the same time or select the same contention slot. Therefore, backoff mechanism is used to alleviate the congestion and reduce the probability of collision by increasing the contention duration. This increase is controlled by an exponent and takes the number of consecutive collisions into account. Hence, using different backoff exponents for different traffic classes can also be considered as a technique for service differentiation as in [58].

- **Transmission slot scheduling**: Reservation-based medium access schemes divide the time into small portions called slot. Although there are plenty of slot assignment techniques in the literature, specific slot assignment methods can be derived according to the requirements of the application. For example, reserving consecutive slots for a video sensor node which transmits delay sensitive real-time video frames can increase the service quality considerably.

- **Changing active time**: MAC protocols employing sleep-listen schedule for energy saving can set the active time of the sensor nodes according to their priority level [49,50]. For example: sensor nodes processing best-effort data may work with 1% duty cycle while nodes processing real-time data are working with 50%. Eventually, lower latency and packet dropping ratio and higher throughput can be achieved for higher priority traffic.

- **Changing adaptation speeds**: Some protocols dynamically adapt themselves to the current network conditions by changing some parameters like CW size or backoff exponent during operation of the sensor node. Using different coefficients for the adaptation of parameters can control the speed of convergence to local optimums, hence can provide service differentiation [49]. Setting smaller coefficients for low priority traffic and bigger coefficients for high priority traffic in case of down-scale adaptation of CW size might be a good example.

- **Changing error correction strength**: MAC protocols utilizing error control mechanisms to provide QoS support can accommodate service differentiation by changing either persistency of retransmissions [60] or strength of the error control codes as mentioned in Section 5.2. Error resiliency of the traffic belonging to different priority classes can be controlled easily and hence, desired level of reliability can be assured for each traffic class.

- **Changing DoA**: As mentioned in Section 5.3, higher DoA needs accumulation of packets at the buffer of the router node which causes longer delays. On the other hand, lower DoA decreases the quality of redundancy elimination and increases the energy consumption. Therefore, employing variable DoA for each traffic class can be a
6. QoS-aware MAC protocols for WSNs

As emphasized in the previous sections, MAC layer of the architecture stack plays a key role in QoS provisioning. There are numerous WSN MAC protocols in the literature [10,11] but few of them take QoS support into account. Since sensor nodes are battery-powered devices, the main motivation of the almost all of the proposed MAC protocols is energy-awareness. However, there is an increasing necessity for efficient QoS-aware MAC protocols parallel to the increasing application fields such as health care, surveillance and process control. In this section, QoS-aware MAC protocols in the literature will be surveyed along with their advantages and disadvantages. We will start with the protocols employing service differentiation and continue with application specific ones. Then, protocols providing indirect support to QoS provisioning will be mentioned. Comparison and classification of the existing QoS-aware MAC protocols for WSNs will conclude this section.

6.1. Protocols with differentiated services

6.1.1. PSIFT

PSIFT [53] is a QoS-aware MAC protocol designed for event-driven applications and it is based on the SIFT protocol [62], which exploits the spatial correlation property of WSNs. SIFT assumes that the first $R$ of $N$ reports of a detected event are the most important part of the messaging and have to be relayed with low latency. $R$ reports will be sufficient for the sink node to accurately identify the event and elimination of redundancy decreases both probability of collision and latency. Authors proposed two methods “Explicit ACK” and “Implicit ACK” for suppressing the unnecessary redundant reports by utilizing the broadcast nature of the wireless medium.

PSIFT is a Carrier Sense Multiple Access (CSMA)-based MAC protocol and provides traffic differentiation by varying the inter frame space (IFS) and contention window (CW) size for each traffic class, as shown in Fig. 2. Traffic classes are prioritized in a dynamic manner based on the traversed number of hops, i.e. the higher number of hops traversed, the higher level of priority that a packet has.

Advantages and disadvantages: Although PSIFT might be a sensible choice for event-driven applications, it is nearly impossible to be used in any other type of applications. Besides, removal of redundancy may result in unreliable data delivery since identification of reports belonging to separate events will be an issue to be solved. Report suppression mechanism decreases the traffic load in the network and leads to mostly idle sensor nodes. This advantage of the PSIFT must be utilized to decrease the energy consumption of the network by integrating a sort of sleep-listen schedule.

6.1.2. Saxena et al. MAC

Saxena et al. MAC [49] aims to offer QoS for multimedia transmission over WSNs and to conserve energy without violating QoS-constraints. This protocol uses a CSMA/CA approach and assumes three types of traffic carried in the network: streaming video, non-real-time and best effort. Basically, the MAC scheme periodically monitors the dynamics of the sensor nodes and the medium, and collects relevant network statistics like transmission failures and transmitted traffic type. Accordingly, the protocol updates the CW size and duty cycle adaptively, based on the gathered information.

Energy conservation is achieved by employing adaptive duty cycles according to the dominantly processed traffic in the sensor node. Hence, each sensor node follows its own sleep-listen schedule. Service differentiation between traffic classes is achieved by using different coefficients for each traffic class to control increase and decrease speed of the CW sizes. Consequently, CW size for higher priority traffic decreases faster than the lower priority where an increase is performed more slowly.

Advantages and disadvantages: Although highly dynamic operation of the protocol adapts well to the changing network conditions, it introduces a significant overhead and complexity. Additionally, idle listening and early sleeping problems most likely to occur since there is no local or global synchronization between sensor nodes. The protocol causes lower-priority packets to suffer from high latencies.

6.1.3. PR-MAC

PR-MAC [52] gives different priorities for each type of event monitored by the sensor nodes and provides service differentiation among these events by varying both CW size and IFS for each of them. The sender node transmits a short pulse to reserve the medium rather than using RTS-CTS exchange. Hence, collisions can only occur during transmission of the burst pulse among nodes of equal priority.

Acknowledgement mechanism is achieved by sending powerful broadcast signals from sink to every node in the network. Moreover, acknowledgement by the intermediate nodes is not implemented. Thus, there is no retransmission scheme in PR-MAC since authors care about the delivery latency of the sensed event more than its reliability.

Advantages and disadvantages: Sink-to-source acknowledgement mechanism requires a very powerful sink node to be heard by every sensor node and seems to be impractical. Also, lack of acknowledgement between relaying nodes disrupts the reliability of the protocol seriously. PR-MAC reserves the medium without RTS-CTS message exchange, and hence reduces the control overhead. However, it may face some problems to support variable size
packet delivery since RTS packets includes the medium reservation duration.

6.1.4. RL-MAC

RL-MAC [50] is a QoS-aware reinforcement learning (RL) based MAC protocol and uses a CSMA scheme. It adaptively changes the duty cycle of the sensor nodes based on not only local observations but also by the observations of neighbor nodes. As a local observation, the number of successfully transmitted and received packets during the active time period is recorded to be used in the duty cycle adaptation with proportional load of the queues. For neighbor observation, a field is added to the packet header to provide information to the receiving node regarding the number of failed transmission attempts by the sender. With this field, RL-MAC tries to save energy while minimizing the number of missed packets due to early sleeping. Traffic load in the network is divided into three traffic categories and service differentiation between them is implemented by varying the CW size of each category.

Advantages and disadvantages: Relatively complex RL based algorithm adapts the network conditions very well but it might not be feasible to be implemented on energy and processing power constrained sensor nodes.

6.1.5. Q-MAC

Q-MAC [54] utilizes intra-node scheduling to select the next serviced packet from five different priority queues and inter-node scheduling to coordinate the medium access among multiple neighboring nodes as seen in Fig. 3. The priority of an incoming packet is determined by two factors. Application layer perspective gives priorities based on the content of the packet and MAC layer does based on traversed hop count. In this way, packets are mapped into predefined five different priority queues including one instant queue that any packet in this queue is served immediately. Within the context of intra-node scheduling, MAX-MIN fairness algorithm [63] is used to control the rate and packetized Generalized Processor Sharing [64] algorithm is used to select the next transmitted packet. For inter-node scheduling, a novel protocol named Loosely Prioritized Random Access (LPRA) is proposed for coordinating the transmission urgencies of the nodes which have packets to send. There are four factors determining the transmission urgency of a node: packet criticality from application point of view, traversed hop count of the packet, remaining energy of the sensor node and queue’s proportional load.

A frame represents single RTS-CTS-DATA-ACK packet exchange and consists of contention period (CP) and transmission period (TP). CP is divided into five smaller contention portions which are exclusive to sensor nodes that have certain level of transmission urgency. As congestion control mechanisms, doubling the CW size is proposed for decreasing the probability of collision and decreasing the packet deadline for alleviating the traffic load. For energy efficiency, sensor nodes follow sleep-listen schedules with fixed duty cycles.

Advantages and disadvantages: Dynamic priority assignment provides robustness against changing conditions of the sensor network. However, calculation of the transmission urgency of a node is relatively complex. Integration of the increasing geometric probability for CW selecting may decrease the collision rate but also may result in higher latencies.

6.1.6. PQ-MAC

PQ-MAC [57] aims to use advantageous features of both contention based and schedule based approaches and uses a hybrid scheme for medium sharing. Global clock synchronization, neighbor discovery and accordingly slot assignment are done during the setup phase and followed by the transmission phase where the real data delivery takes place.

The slot assignment within the setup phase considers the two-hop distance neighbor nodes and allocates different time slots based on the DRAND [65] algorithm and the frame size is determined by the time frame rule of the Z-MAC [66] protocol. Owner node of a specific transmission slot, assigned in the setup phase, has an exclusive right to send the data in it. If the owner of the slot does not have any data to send or has lower priority data, non-owners of the slot can contend for the slot based on priorities of their data.

The Super Frame (SF) structure of the PQ-MAC consists of two sub frames: Data Frame (DF) which is used for data delivery and Control Frame (CF) which used for the sleep-listen schedule. An adaptive sleep-listen schedule is used for energy efficiency and synchronization between neighboring sensor nodes is provided by generating sequence of bits indicating whether the sensor node will sleep or be awake during the corresponding time slot. In Fig. 4, the medium access prioritization mechanism is presented for three different traffic classes. Only the owner of the slot can access privileged contention windows T0, T2 and T4 while non-owners can contend during T1, T3 or T5 with respect to their traffic types.

Advantages and disadvantages: The neighborhood of the sensor nodes, relay nodes or cluster heads may change frequently because of the dynamic nature of the WSNs, as mentioned earlier. Therefore, accuracy of the slot assignment performed once at the beginning of the setup phase will be obsolete during the transmission period gradually. In heavy traffic conditions, PQ-MAC behaves like a TDMA based protocol since almost all nodes will have a packet to send and use its own transmission slot. This improves the channel utilization and reduces the
probability of collision significantly at the cost of tight clock synchronization.

6.1.7. QoMOR

A QoS-aware MAC protocol using Optimal Retransmission (QoMOR) [60] is designed for the intra-vehicular sensor networks and assumes the sensor nodes have only the transmission capability. Since sensor nodes cannot receive any acknowledgement from the sink node or detect collisions, authors derived an optimization problem to find the minimum number of retransmissions required to achieve a certain level of frame delivery probability bounded by a maximum delay threshold.

Theoretical analysis of the single QoS class is presented based on the derived optimization problem and it is extended to multiple QoS classes where each sensor node is a member of a QoS class. An algorithm is also given for the two QoS classes case to approximate the optimum number of retransmissions for guaranteed frame delivery probability.

Advantages and disadvantages: Reduction of receiver hardware decreases the cost of the sensor nodes considerably. One way transmission of the data and absence of coordination makes QoMOR very lightweight and simple solution for one-hop sensor networks. However, as authors indicated, it is very hard to achieve an acceptable level of frame delivery probability with stringent delay constraints under dense networks and this objective becomes more challenging as the frame size increases.

6.1.8. IEEE 802.15.3/802.15.4 and extensions

Besides discussing the QoS-aware MAC protocols designed for WSNs, in this section we discuss the state of the art in related MAC layer standards. The aim of IEEE 802.15.3 [68] standard is to develop an ad hoc MAC layer for high data rate wireless personal area networks (WPANs) and a physical layer that can reach up to 20Mbps. The standard is geared towards handling voice, images and file transfers and it has an operational transmission range of approximately 10 m. Basically, the standard is specified for higher data rate scenarios and does not address the requirement of energy efficiency or other QoS requirements in WSNs.

The IEEE 802.15.4 standard [69,68], which is used as a basis for the ZigBee, WirelessHART, and MiWi specifications, has been originally designed for low-rate WPANs. The standard is then adopted by WSNs, interactive toys, smart badges, remote controls and home automation, operating on license-free ISM bands. IEEE 802.15.4 is intended as a specification for low-cost, low-powered networks with no critical concerns about throughput and latency. Therefore, QoS issues have not been the main concern in the original specification. Later, the IEEE 802.15.4a Task Group was created with the goal of defining a new physical layer, which is able to provide higher data rates and high-accuracy ranging capabilities. New releases of the standard focus on using UltraWide Band (UWB) and chirp signals as alternative physical layer technologies to overcome the bandwidth limitations. UWB can achieve bit rates varying approximately between 0.1 Mbps and 26 Mbps. However, besides higher data rates, other QoS issues, such as latency, reliability, are not addressed in the specification. Instead, there exists a number of studies to improve the performance of IEEE 802.15.4 MAC standard in terms of QoS support [58,70–72]. Since they mainly adopt similar strategies, we believe that, surveying one of the examples will be sufficient to understand the basics of QoS support in 802.15.4 MAC.

In [58], authors derived an extension for IEEE 802.15.4, beacon enabled slotted CSMA-CA standard to provide service differentiation among sensor nodes based on their application-specific level of importance. Two mechanisms are proposed to realize service differentiation: variable contention window size and variable backoff exponent. A mathematical model based on the discrete-time Markov chain is also presented to evaluate the throughput, delay and packet drop probability performance of the modified 802.15.4 standard.

Advantages and disadvantages: Since it is a service differentiation add-on scheme proposed for a well-known MAC protocol, it can be widely used in all IEEE 802.15.4 compatible sensor devices. However, priorities can only be assigned to sensor nodes statically beforehand which makes this proposal inappropriate for multi-modal sensor networks.

6.1.9. I-MAC

I-MAC [56] uses a hybrid TDMA/CSMA scheme for medium access and basically introduces a prioritization mechanism for Z-MAC [66]. There are two phases during execution as in Z-MAC: set-up phase in which neighbor discovery, slot assignment, local framing and global synchronization occurs; and transmission phase where time is divided into slots.

There are three predefined priority levels mapped to each sensor node according to its role in the network.
I-MAC anticipates dynamic prioritization where sensor nodes set their own priority level according to their local observations like traffic load, remaining energy and distance to sink. Authors propose a scheduling algorithm called DNIB [73] and time slots are assigned to each sensor node based on this algorithm. Owner of the time slot has guaranteed access in that particular slot and this guarantee is provided by employing Arbitration Interframe Space (AIFS) for non-owner sensor nodes. If the owner has no data to send or the slot is not owned, non-owners can compete for transmission. Service differentiation among non-owners is provided by adopting different CW sizes for each priority level.

Advantages and disadvantages: Although I-MAC combines the strength of both TDMA and CSMA schemes, it still needs tight clock synchronization which is a well-known drawback of TDMA schemes. Authors developed a novel scheduling algorithm and achieved better utilization than of Z-MAC. However, possessing up-to-date neighbor information and slot schedule in highly dynamic sensor networks is a major challenge.

6.1.10. Diff-MAC

Diff-MAC [74] is a CSMA/CA based QoS-aware MAC protocol with differentiated services and hybrid prioritization. Diff-MAC aims to increase the utilization of the channel with effective service differentiation mechanisms while providing fair and fast delivery of the data. Primary application field of the Diff-MAC is wireless multimedia sensor networks which commonly carry QoS-constrained heterogeneous traffic.

Diff-MAC has some key features to provide QoS: (i) Fragmentation and message passing feature fragments the long video frames into smaller video packets and transmits them as a burst which in turn reduces the retransmission cost in case of MAC failures. (ii) Diff-MAC can adjust its CW size according to the traffic requirements to reduce the number of collisions and keep the packet latencies as small as possible. (iii) Diff-MAC adapts duty cycle of the sensor nodes according to dominating traffic class and tries to balance both energy consumption and delay. (iv) Intra-node and intra-queue prioritization feature provide fair delivery of the data among all sensor nodes and among all traffic classes respectively to avoid intolerable performance.

Advantages and disadvantages: Fast adaptivity to changing network conditions and network-wide fairness of Diff-MAC make it a very strong candidate for multimedia sensor applications. However, monitoring network statistics and dynamic adaptation are complex and overwhelming operations. Additionally, although lack of sleep-listen synchronization between neighboring sensor nodes improves the protocol scalability, it also increases the packet latencies caused by early sleeping.

6.1.11. SASW-CR

SASW-CR [51] is a slotted Aloha based MAC protocol for Ultra-wideband (UWB) sensor networks with QoS support. Authors assume all nodes in the network are classified as high or low priority depending on the traffic they generate and service differentiation between them is achieved by using disjoint contention windows. A cooperative retransmission technique based on overhearing is also utilized to provide fast and reliable data delivery.

Each sensor node maintains two queues; namely data queue which stores the created data packets by the sensor node itself and overhearing queue which stores overhead packets during transmission belonging to neighboring sensor nodes. Sensor node may transmit a packet either from its data queue or overhearing queue depending on its mode. In selfish mode, a node always transmits its own packet first while in selfless mode, node selects a packet from the overhearing queue.

In Fig. 5, a high priority sensor node (HP) which tries to send two data packets ($P_1, P_2$) to the sink is depicted. Since $P_1$ could not be relayed by its creator, transmission of $P_1$ is completed by overhearing low priority sensor node (LP) where $P_2$ is directly sent to the sink. In this way, SASW-CR decreases the packet latencies and alleviates the link failure effects.

Advantages and disadvantages: Although cooperative retransmission improves the MAC layer performance, each node must acquire acknowledgements broadcast by the sink node in order to eliminate unnecessary copies of overheard packets. Moreover, maintaining such a mechanism requires continuously active sensor nodes which results in high energy consumption.

6.1.12. EQ-MAC

EQ-MAC [75] is designed to provide QoS support for cluster based single-hop sensor networks by service differentiation and uses a hybrid medium access scheme. The protocol is composed of two parts: Classifier MAC (C-MAC) and Channel Access MAC (CA-MAC).

C-MAC classifies the received data packets into four priority levels according to the importance of the packet assigned by the application layer and uses a queuing architecture similar to Q-MAC [54]. This architecture includes an instant queue and packets stored in that queue are served immediately.

CA-MAC is responsible for medium sharing and consists of four phases repeated in each frame: Synchronization, Request, Receive Scheduling and Data Transfer. During Synchronization, Request and Receive Scheduling phases; sensor nodes get synchronized, contend to send their channel requests to cluster head and receive scheduling messages broadcast from cluster head. Only control messages are exchanged in these first three phases and medium is

![Fig. 5. Cooperative retransmission in SASW-CR [51].](attachment:image.png)
shared based on CSMA/CA. In the last phase of CA-MAC which is Data Transfer, each sensor node follows the transmission schedule received from cluster head and accesses to the medium without contention. Sensor nodes that have no data to send or could not manage to acquire a transmission slot go to sleep state during this phase for power saving.

**Advantages and disadvantages:** Probability of collisions and energy consumption are reduced by using contention based medium access for short periodic control messages and by scheduled medium access for long data packets. However, authors try to overcome classical synchronization problem of TDMA scheme by employing a SYNC phase at the beginning of each frame which brings an extra overhead to the protocol. Moreover, EQ-MAC is designed for single-hop cluster based sensor networks and cannot handle multi-hop transmissions. Also, clustering algorithm is not included in the MAC protocol itself.

### 6.2. Application-specific protocols

#### 6.2.1. EQoSA

EQoSA [76] is a hybrid MAC protocol which is designed to provide QoS support especially for video and image transmission over sensor networks. Basically, EQoSA modifies the fixed session size of the BMA [77] protocol and uses dynamic session sizes regarding the number of active sensor nodes and their traffic loads. During the contention period, each node reports whether it has data to transmit or not. The cluster head then performs the slot assignment and broadcasts to all sensor nodes. In this way, EQoSA accommodates bursty traffic by allocating the required number of data slots for each sensor node in each session.

**Advantages and disadvantages:** EQoSA suffers from the traditional time synchronization problem of TDMA based schemes and only has the ability to accommodate bursty traffic load rather than a proper service differentiation mechanism. Moreover, it needs more powerful cluster heads within the sensor network to perform and announce the slot assignment.

#### 6.2.2. Suriyachai et al. MAC

Suriyachai et al. MAC [78] provides QoS support by giving deterministic bounds for node-to-node delay and reliability, hence can be a suitable candidate for applications requiring absolute delay and reliability assurance. Authors employed a collision-free TDMA scheme and divided the time axis into fixed-length portions called epochs. In each epoch, a sensor node has \( k \) exclusive slots for only single DATA-ACK message exchange. All of \( k \) slots are used for retransmission until a successful packet delivery occurs by receiving an ACK message. Accordingly, node-to-node delay is bounded by the duration of an epoch theoretically. If a sensor node does not have any data to send, it sends a simple control message at the first reserved slot indicating that it will not send anything in this epoch.

\( K \) retransmission slots are distributed in the epoch so as to obtain maximum temporal distance for mitigating the burst errors in the wireless channel. By assuming independent bit error rates, they also give a guaranteed theoretical bound for reliability. Energy consumption is reduced by employing different duty cycles for each sensor node depending on their number of child nodes in the predetermined data gathering tree.

**Advantages and disadvantages:** Since each node synchronizes its clock with its parent node, synchronization errors can propagate increasingly. Also, each node must be aware of its position in the data gathering tree for slot assignment and duty cycling. Therefore, Suriyachai et al. MAC does not scale well for large networks. Moreover, although it can bound delay and reliability, it is impossible to obtain proper throughput performance by reserving whole epoch for only single data transfer.

### 6.3. Protocols with indirect QoS support

Although the previous sections summarize a variety of QoS-aware MAC protocols in the literature, there are still some other protocols that we need to mention. These protocols support QoS provisioning even though they are not designed to provide it as a primary objective. Most of these indirect-QoS-aware MAC protocols adapt themselves to the current network conditions and achieve better performance in QoS aspect.

WiseMAC [79] tries to reduce the energy consumption by determining the length of the preamble dynamically. CA-MAC [80] adapts the duty cycle of the sensor nodes based on their buffer load and priority of the packets stored in the buffer where TRAMA [81] adapts the number of time slots reserved for each sensor node according to their current traffic rate. I-EDF [55] is a MAC protocol based on earliest-deadline-first and tries to provide latency requirements of delay-bounded data. LWT-MAC [82] responds effectively to sporadic changes in the event-based sensor networks by switching to unscheduled medium access under low traffic load and to scheduled medium access under high traffic load. Jiang et al. [83] propose a fuzzy algorithm which aims to reduce the packet error rate and prolong the network lifetime by adjusting the transmission power of the sensor nodes adaptively.

Some protocols modified the S-MAC protocol [84], which is a well-known sensor MAC protocol, and proposed dynamic versions of it. T-MAC [85] adapts the active time in S-MAC while DSMAC [86] adds a dynamic duty cycle feature to S-MAC. TA-MAC [87] modifies the static CW mechanism of S-MAC and adapts itself to the current traffic load. PSMAC [88] is a joint MAC and physical layer protocol and introduces a transmission power control mechanism to S-MAC.

Although Lump [61] protocol operates between the link and the network layer, it can be considered as a MAC component rather than a complete MAC protocol. Lump utilizes a differentiated data aggregation technique to provide QoS support. The aim of the protocol is to reduce radio communication and minimize energy consumption while fulfilling the specific latency requirements of each traffic type.

As mentioned in Section 5.5, clustering is a viable technique for QoS provisioning. QBCCDP [89] supports video and image transmission with dynamic clustering and provides QoS support in terms of delay and bandwidth. QUATRO [90] proposes collaboration of MAC and network
layers. It utilizes clustering and scheduled medium access to achieve QoS provisioning.

Although there exist studies on addressing QoS challenges at the routing layer [91,92,27], to the best of our knowledge, there is no QoS-aware MAC protocol designed for mobile wireless sensor networks. However, there exist a few MAC protocols that address the challenge of mobility with adaptive MAC protocols. For instance, in [93], the TDMA-based LMAC protocol [94], which is designed for static WSNs, is modified for mobility support. Different than the static LMAC protocol, in adaptive LMAC, nodes update their selected time slots for medium access whenever the neighborhood of a node changes due to mobility. However, adaptation of the frame lengths or number of time slots per frame for QoS support are not discussed or implemented in this study. In another study [95], authors propose a modification of the S-MAC protocol [84], called MS-MAC for mobile WSNs. In MS-MAC, nodes discover the presence of mobility within their neighborhood based on the received signal levels of periodical SYNC messages from the neighbors. If a node detects a change in the strength of a signal received from a neighbor, it concludes that the neighbor or the node itself are moving and adaptively changes the schedule that it is following according to the mobility patterns in the neighborhood. MOBMAC [96], is another adaptive MAC protocol designed for mobile WSNs. MOBMAC addresses the problem of frame losses caused by the communication signals experiencing mobility induced effects such as Doppler Shifts. It introduces an adaptive frame size predictor, using an Extended Kalman Filter to predict an optimal frame size for every transmission. A smaller frame size is predicted when the signal characteristics are poor (i.e. when the signal is Doppler shifted) and larger frame sizes are predicted when the quality of the channel improves (i.e. when the nodes are stationary). By transmitting a small frame size in a bad channel, MOBMAC reduces the transmission power since smaller frames need lower transmission power compared to larger frames and also reduces the probability of error occurrence since it is less in a smaller frame than that of a large frame.

Since underwater and underground sensor networks are not as practical and mature as traditional terrestrial WSNs, there is no noticeable QoS aware MAC protocol in the literature for UW-ASNs and WUSNs. However, there is a group of MAC layer proposals for underwater sensor networks which can be further improved for QoS support. UW-MAC [97] is a CDMA based MAC protocol for underwater sensor networks and it has three objectives, which are high throughput, low delay and low energy expenditure. UW-MAC achieves these objectives easily in deep waters while it tries to adaptively find the optimal tradeoff among the objectives in shallow waters. In [98], authors propose UWAN-MAC for stationary underwater sensor networks that have to operate under long, unknown propagation delays and they select energy efficiency as their main performance metric. UWAN-MAC uses a CSMA based scheme and employs sleep/listen schedules to conserve energy. Local synchronization among neighboring sensor nodes is achieved by means of periodic SYNC messages. Although there are other MAC protocols proposed for ad hoc underwater acoustic networks such as [99–101], we will not go into details of them in this work. However, reader may refer to [102] for an overview of networking protocols for underwater wireless communications.

Some researchers approach QoS provisioning with a wider perspective and propose frameworks or architectures rather than constraining the problem to a single communication layer. RAP [103] is a real-time communication architecture for large-scale WSNs and introduces Velocity Monotonic Scheduling which forwards the packets to their destinations at requested velocity, hence tries to accurately fulfill the end-to-end deadline requirement of real-time traffic. Yuan et al. [104] proposed an integrated single framework to jointly optimize the energy efficiency and QoS. Fallahi and Hossain [105] derived a dynamic power management framework for wireless video sensor networks to achieve energy saving while providing QoS support. Troubleyn et al. proposed AMoQoSA [106], which is an adaptive modular QoS architecture for heterogeneous sensor networks. Aim of this architecture is to continuously deliver QoS support by activating a set of QoS techniques according to capabilities of the sensor nodes.

6.4. Comparisons

In Table 2, we summarize the general aspects of the QoS-aware MAC protocols that we have discussed for WSNs. The table also presents comparisons of the discussed algorithms in two groups, namely protocols with differentiated services and application-specific protocols. The Type column shows the type of MAC mechanism(s) used in the protocol. Service Differentiation column specifies whether the protocol supports service differentiation or not whereas the Priority Assignment column presents whether the protocol assigns priorities to different traffic types and if it does whether it is static, dynamic or hybrid. The Synchronization field shows whether the protocol requires synchronization or not. The Energy-awareness column is important to show whether the protocol provides energy-awareness together with QoS provisioning which are known to be conflicting requirements. The Complexity column demonstrates the level of complexity in the execution of the protocol. Finally, the Scalability field shows how scalable the protocol is with the increased number of sensor nodes and complexity within a WSN.

Most of the protocols provide random medium access, i.e., CSMA, or propose hybrid solutions such as CSMA and TDMA. We observe that instead of providing deterministic QoS guarantees, majority of the protocols follow a service differentiation approach by classifying data packets according to their type and associated network parameters at the MAC layer are tuned according to the requirements of different types. Those protocols that provide service differentiation usually assign static priorities to the traffic types since this is simpler to manage. However, dynamic priority assignment may be necessary where the priorities of different packet types may vary in time. Traffic adaptivity is usually not supported whereas most of the protocols do not require synchronization since they allow for random access. As the MAC protocols provide QoS provisioning, their complexity increases but still they should be
processed by the sensor devices without any resource problems. Although WSNs are becoming popular among complex applications that require fast and efficient data delivery, energy awareness is still a major requirement and most of the protocols support energy efficient communication in the network. In terms of scalability, we observe both trends: good and weak protocols in terms of scalability but one should not forget that WSNs are composed of hundreds, thousands and even more devices and the protocols should be able to work with these numbers of nodes.

As we mentioned before, performance of the MAC protocols for WSNs are highly application dependent. Therefore, we need to evaluate the performance of all surveyed protocols under the same application or simulation environment, which is quite hard to be done, in order to make accurate quantitative comparisons in terms of communication delay, delay jitter, throughput, energy efficiency, lifetime, etc. However, for those interested, please check the individual papers of these protocols for small scale qualitative comparisons between their competitors. For instance in [74] we compare the performance of DiffMAC with Saxena et al. MAC [49] and give qualitative results in terms of lifetime, delay, energy efficiency and delivery rate.

Additionally, a classification of these protocols is provided in Fig. 6. As mentioned, we observe two main trends in QoS-aware MAC protocols for WSNs: protocols that follow differentiated services approach and protocols that provide application specific QoS support. Protocols that provide service differentiation can further be classified as the protocols that provide static differentiation (i.e., static parameters are tuned at the MAC layer), protocols with dynamic differentiation where dynamic parameters are tuned at the MAC layer, such as the remaining time till the packet deadline, and the protocols with hybrid QoS support where both static and dynamic parameters are taken into account as discussed in Section 5.6. Among the protocols that we have surveyed in Section 6, [49–52,57,60,58,75] provide static differentiation whereas

### Table 2
Comparison of QoS-aware WSN MAC protocols in the literature.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Type</th>
<th>Service diff.</th>
<th>Priority assignment</th>
<th>Traffic adaptivity</th>
<th>Synchron.</th>
<th>Energy awareness</th>
<th>Complexity</th>
<th>Scalability</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSIFT [53]</td>
<td>CSMA</td>
<td>Dynamic</td>
<td>No</td>
<td>None</td>
<td>No</td>
<td>Low</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Saxena et al. [49]</td>
<td>CSMA</td>
<td>Static</td>
<td>Yes</td>
<td>None</td>
<td>Yes</td>
<td>High</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>PR-MAC [52]</td>
<td>CSMA</td>
<td>Static</td>
<td>No</td>
<td>None</td>
<td>Yes</td>
<td>Low</td>
<td>Weak</td>
<td></td>
</tr>
<tr>
<td>RL-MAC [50]</td>
<td>CSMA</td>
<td>Static</td>
<td>Yes</td>
<td>None</td>
<td>Yes</td>
<td>High</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Q-MAC [54]</td>
<td>CSMA</td>
<td>Hybrid</td>
<td>No</td>
<td>None</td>
<td>No</td>
<td>Moderate</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>PQ-MAC [57]</td>
<td>TDMA/CSMA</td>
<td>Static</td>
<td>No</td>
<td>Locally</td>
<td>Yes</td>
<td>High</td>
<td>Weak</td>
<td></td>
</tr>
<tr>
<td>QoMOR [60]</td>
<td>ALOHA</td>
<td>Static</td>
<td>No</td>
<td>None</td>
<td>No</td>
<td>Low</td>
<td>Weak</td>
<td></td>
</tr>
<tr>
<td>802.15.4 excl. [58]</td>
<td>CSMA</td>
<td>Static</td>
<td>No</td>
<td>None</td>
<td>No</td>
<td>Moderate</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>L-MAC [56]</td>
<td>TDMA/CSMA</td>
<td>Dynamic</td>
<td>Yes</td>
<td>Locally</td>
<td>No</td>
<td>Moderate</td>
<td>Weak</td>
<td></td>
</tr>
<tr>
<td>Diff-MAC [74]</td>
<td>CSMA</td>
<td>Hybrid</td>
<td>Yes</td>
<td>None</td>
<td>Yes</td>
<td>High</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>SAW-S-CR [51]</td>
<td>ALOHA</td>
<td>Static</td>
<td>No</td>
<td>None</td>
<td>No</td>
<td>Moderate</td>
<td>Weak</td>
<td></td>
</tr>
<tr>
<td>EQ-MAC [75]</td>
<td>TDMA/CSMA</td>
<td>Static</td>
<td>Yes</td>
<td>Locally</td>
<td>Yes</td>
<td>High</td>
<td>Weak</td>
<td></td>
</tr>
<tr>
<td>EqSA [76]</td>
<td>TDMA/CMDA</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Locally</td>
<td>Yes</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Suruya et al. [78]</td>
<td>TDMA</td>
<td>-</td>
<td>No</td>
<td>Network-wide</td>
<td>Yes</td>
<td>Moderate</td>
<td>Weak</td>
<td></td>
</tr>
</tbody>
</table>

### Fig. 6. Classification of QoS-aware MAC protocols.
protocols [53,56] provide dynamic differentiation and [54,74] propose a hybrid approach.

7. MAC layer design tradeoffs for QoS provisioning

Critical decisions must be taken during the design phase of the protocols. These design tradeoffs need to be studied extensively and must be chosen according to specific requirements of the sensory application since they will provide a basis for the protocol. In this section, we will evaluate MAC layer design tradeoffs and highlight their advantages and disadvantages from the QoS point of view. Most of the design tradeoffs are related with service differentiation since it is an integral part of the QoS provisioning and majority of the MAC layer protocols provide differentiated services.

7.1. CSMA vs. TDMA schemes

TDMA scheme divides the time into smaller slots and sensor nodes communicate within their own slots in a contention-free manner. Hence, a centralized or distributed slot assignment algorithm is needed in TDMA to decide which sensor node will transmit its packet in which transmission slot. As a result of this scheduling, wireless channel can be utilized well. Moreover, theoretical QoS bounds such as throughput and latency can be given since each sensor node knows when to transmit. This also brings the ability to easily adopt a sleep-listen schedule for energy saving. However, the scheduling algorithm must have information regarding the number of sensor nodes and their positions in order to make a proper slot assignment. Although some examples of scheduling algorithms require only the information of neighboring sensor nodes, they still require a neighbor discovery operation.

Having the topological information of the network or neighbor discovery is not sufficient for slot assignment in the long term. Depletion of energy resources, hardware malfunctioning, node mobility, link failures can cause frequent topology changes in WSNs and up to date state of the network must be obtained periodically for accurate slot assignment. Thus, TDMA does not scale well as the size of the network increases. Even accomplishing slot assignment is not enough to properly operate TDMA, tight clock synchronization between sensor nodes is still needed in order to prevent transmission slot violations originated from clock drifts. Besides, contention-free approaches are less likely to be able to respond well in case of variable and bursty traffic conditions and might cause intolerable performances.

On the other hand, contention-based schemes where sensor nodes contend to access the shared medium are very easy to implement and more appropriate for infrastructure less sensor networks. CSMA scheme does not require any additional information related with the network topology or offered traffic load. Thus, performance of the CSMA schemes are not as dependent as TDMA schemes on the network topology and scales well for changing network size and density. Moreover, contention-based schemes can handle bursty and sporadic traffic since sensor nodes do not have to follow a transmission schedule. However, collisions might occur in contention-based schemes with an increasing probability as the contender nodes or offered traffic load increases and this causes extra delivery latency, high energy expenditure and retransmissions. Hence, they cannot guarantee a certain level of service quality. Although some medium reservation mechanisms are proposed to avoid collisions like RTS/CTS, they introduce some overhead. Thus, efficient reservation, contention and back-off strategies must be employed.\footnote{Protocols using adaptation coefficients can also be classified as adaptive.}

Yet another medium sharing scheme, called hybrid scheme, developed to overcome drawbacks of both scheduled and unscheduled methods. Hybrid schemes can classify the packets (e.g. data, control, low priority, high priority) and choose the proper way to access the medium regarding the belonging class of that particular packet. Another method is to melt these two techniques into one by letting the non owner sensor nodes of a previously assigned TDMA time slot to contend for transmission chance. Good combination of existing techniques can utilize the network resources and provide significant energy saving which in turn has to deal with disadvantages of the each composing technique.

7.2. Static vs. dynamic priority assignment

Selected priority assignment method is quite important for QoS support since resource sharing among different priority classes is carried out according to their importance. Priorities can be assigned to the sensor nodes as well as to the packets created by them. Assigning the priorities statically is not a complex issue since there is no need for any observation or calculation. Once the priority is given, it does not change during the operation of the sensor node or delivery of the packet. On the other hand, dynamic priority assignment needs some additional assessments and priority reassignment accordingly in every triggering event (e.g. arriving another hop for packets, role changes for sensor nodes) which brings an extra overhead to the QoS mechanism. However, adaptive changes regarding the importance of the packet or the sensor node can significantly improve the performance of the QoS mechanism.

Decision parameters needed in the dynamic priority assignment may not be present in the format of the packet so that additional fields in the packet format are required. This causes bigger packets which means longer transmission times and energy consumption. It should be sufficient to have a simple priority field in the header of the packet in the case of static prioritization. Moreover, the dynamic priority assignment method mostly requires decision parameters (mentioned in Section 5.6.1) which are not MAC-specific and necessitate cross-layer mechanisms.

7.3. Single-queue vs. multi-queue architecture

Protocols that employ differentiated services classify the carried traffic into different priority levels and the
MAC protocol maintains either a single queue for every traffic type or separate queues for each of them. Main drawback of the single-queue scheme is the high cost of managing relatively long data queue. Since different priority packets are stored in the same queue, it is impractical to keep them sorted and process the packets according to their priorities. On the other hand, the multi-queue scheme chops the long single queue into pieces and employs smaller different priority queues. In this way, packets can be served with a simple FIFO fashion for each priority and additional sorting or searching operations are not needed anymore. However, multi-queue systems have to sacrifice the accuracy of the prioritization if there are more priority levels than the number of available queues since all packets in the same queue are treated as they all have an equal priority. Moreover, in case of multi-queue systems, a fair and QoS-aware packet scheduler must be integrated to select the next serviced queue regarding the requirements of the classified traffic. If not, explicit precedence might cause intolerable performances for lower priority traffic. In case of multi-queue architecture, reader can refer to [107] where a queuing analytical framework for the performance evaluation of MAC protocols with service differentiation is proposed.

7.4. Packet scheduler

In single-queue architectures, there is no need to use a packet scheduler. However, it is mandatory in multi-queue architectures to select the next serviced queue. There exists two design methods for the packet scheduler. The first method is serving the higher priority queue always prior to the lower priority queue explicitly and the second method is utilizing some kind of fair scheduling between the queues of different priority packets.

Main drawback of the explicit prioritization is possibility of intolerable performance for lower priority traffic in terms of latency, successful packet delivery ratio. However, the higher priority traffic achieves relatively better performance since it is always served first. Also, the explicit prioritization can be chosen for the sake of simplicity since it is easy to implement and operate.

There exist many techniques for fair scheduling such as weighted round robin [108], weighted fair queuing [109], deficit round robin [110] to be used in the second method. Integrating a fair scheduling mechanism brings some performance degradation for higher priority traffic since it makes a selection among all nonempty queues. However, a small sacrifice from performance of higher priority traffic results in remarkable performance increase for the lower priority traffic. Also, employing a fair scheduler requires an additional decision phase before each transmission attempt.

8. Properties of a well-designed QoS-aware MAC protocol

As mentioned earlier, the major problem in WSNs is lack of resources. The energy scarcity leads the resource constraints since it will be impossible to use a sensor node anymore with depleted batteries and it becomes totally useless. Therefore, although we are talking about QoS provisioning, first of all, the designed MAC protocol must also be energy efficient. Besides energy, sensor nodes also have limited resources in terms of memory and processing capability. Hence, computationally complex and overwhelming algorithms are not feasible. Moreover, the wireless channel must be well-utilized in order to provide better QoS support since bandwidth scarcity is another challenging issue in WSNs.

The designed QoS-aware MAC protocol must be scalable since WSNs can be composed of excessive number of sensor nodes or deployed to large areas. For this reason, distributed and unscheduled MAC protocols seem to be more suitable to autonomous and ad hoc nature of the WSNs. Moreover, node mobility, environmental effects or node malfunctioning may result in highly dynamic network topologies which makes the adaptive MAC layer requirement a must.

Service differentiation mechanisms can be counted as the most effective way of sharing network resources, especially in resource constrained WSNs. However, integration of service differentiation propounds another issue, which is the necessity for fair and accurate priority assignment methods in order to achieve better QoS performance. Since the poor prioritization of the traffic causes non-utilized network resources, changing network conditions must be taken into account and "dynamic priority assignment" methods must be utilized.

Features listed in this section must exist in a well-designed QoS-aware MAC protocol but not enough to be one. Developers must keep in mind that QoS support in WSNs are highly application-specific. Hence, the performance of the QoS-aware MAC protocols extremely depends on the requirements of the application. For example: delay intolerant real-time applications mostly necessitate fast delivery of the data, while mission critical applications require reliable communication. Therefore, "application-specific requirements" need to be identified with great care and must be used as a primary factor for design tradeoffs.

9. Open issues and future research directions

Application fields of the WSNs are growing rapidly as the capabilities of the tiny sensor devices improve and these applications mostly require varied types of quality assurance. Moreover, diversity of the applications yields to heterogeneous WSNs composed of multimodal sensor nodes which provide more than one functionality by delivering multiple types of traffic. Therefore, novel MAC protocols which have the ability to fulfill the diverse QoS requirements of heterogeneous sensor networks are required.

Heterogeneity of the sensor devices not only introduces challenges but also advantages as well. In recent studies, it is possible to see WSNs composed of several types of sensor devices which have diverse set of capabilities (e.g. energy, communication range, sensing and processing capability). Therefore, envisioned MAC protocols must exploit this diversity in favor of the QoS provisioning by
dynamically adapting themselves to the available resources in the sensor device on which they operate.

When we talk about multimodal WSNs, one certain type is the multimedia WSNs which include cameras and microphone sensors besides scalar sensors. As it is widely studied in other types of wireless networks, delivery of multimedia data has different requirements than the delivery of scalar data, such as higher throughput, bounded delay and image quality. Therefore, novel QoS-aware protocols should be developed to meet the requirements of multimedia WSNs.

With the latest operating systems for WSNs and with the increased popularity, it is possible to have multiple applications running on the same network. This certainly leads to larger amounts of data to be transmitted in the network and handling the traffic, often with different priority levels, in an efficient way becomes a major issue. Protocols to support multiple applications with different QoS requirements running on the same network is another direction of research that should be further investigated. Besides WSNs running multiple applications, different WSNs may coexist in the same spatial domain, i.e. within each other’s neighborhood, and this may cause to share the wireless medium, creating interference and contention on each other. Although different networks may adopt different MAC schemes and QoS provisioning, they need to collaborate and fairly share the wireless resources in the case of co-existence. Therefore collaborative QoS provisioning between coexisting networks may be another topic for further research.

Although we have mainly focused on static WSNs, it is possible to have mobile sensor devices or mobile sink nodes depending on the application requirements. Mobility brings extra challenges in terms of QoS provisioning due to increased dynamics in the network, on top of the ones we have discussed in Section 3. Topology of the network, links between wireless sensor devices change frequently which make it difficult for the QoS-approaches to provide efficient differentiation. In this respect, protocols with dynamic and hybrid differentiation should be adopted and further investigated to meet the requirements of mobile WSNs.

As we briefly mentioned, energy awareness and some QoS requirements, such as high throughput, can be conflicting design factors in WSNs. Theoretical studies that address the tradeoffs between such conflicting requirements could add an important value in terms of providing QoS for WSNs not only at the MAC layer but also for different layers of the protocol stack.

Most of the protocols that we have surveyed in this paper are only evaluated through simulations. However, implementation on real hardware and evaluations on real testbeds would be very useful to avoid the unrealistic assumptions in simulation environment and to evaluate whether the developed protocols meet the resource limitations of real sensor hardware in terms of processing power, memory and energy efficiency. Therefore, implementation of existing and new protocols on real hardware and comparing their performances on testbed environments are other open topics that we identify in the current literature.

According to the comparisons and classification presented in Section 6.4, instead of providing deterministic QoS guarantees, majority of the protocols follow a service differentiation approach and most of the schemes follow either a static differentiation or a dynamic differentiation whereas we could find only one study that focuses on hybrid approach. Hybrid approaches are important since they combine both the static and dynamic parameters to be differentiated at the MAC layer and present a rather extensive solution. Therefore, hybrid service differentiation approaches can be further investigated in future studies to provide efficient service differentiation at the MAC layer.

Since QoS provisioning is not a layer-specific issue and spans all layers in the communication protocol stack, cross-layer mechanisms provide better QoS at the expense of non-modularity by jointly optimizing and merging all layer protocols into single one. Therefore, application-specific cross-layer QoS support mechanisms might be a promising solution for QoS provisioning in resource constrained sensor networks.

10. Conclusions

Current WSNs are not only used for traditional low data-rate applications but also for more complex operations which require efficient, reliable and timely collection of large amounts of data. Moreover, they are not only composed of sensor devices which generate scalar data but also the use of video and microphone sensors are becoming common. Increasing capacities of the sensor nodes, variety of the application fields and multimodal use of sensors require efficient QoS provisioning mechanisms in WSNs. With these requirements in mind, we have focused on the perspectives, challenges, metrics, parameters and requirements of QoS-aware MAC protocols for WSNs in this paper and surveyed the existing protocols together with their comparisons and classifications. According to this survey, we observe that instead of providing deterministic QoS guarantees, majority of the protocols follow a service differentiation approach by classifying data packets according to their type and packets of different types are treated according to their requirements by tuning the associated network parameters at the MAC layer. There are also a few application-specific protocols and protocols that provide indirect QoS support by differentiating the MAC parameters according to the network conditions. Design tradeoffs and open research issues are also investigated to point out the further possible investigations in the field of QoS provisioning in WSNs at MAC layer to contribute to the further research efforts in the field of WSNs.

Acknowledgments

This work is supported by the Scientific and Technological Council of Turkey (TUBITAK) under the Grant No. 108E207, by the Turkish State Planning Organization (DPT) under the TAM Project, number 2007K120610, by the European Community’s Seventh Framework Programme (FP7-ENV-2009-1) under the grant agreement FP7-ENV-244088 “FIRESENSE” and by the Bogazici
References


A. Koubaa, M. Alves, B. Nefzi, Y.Q. Song, Improving the IEEE 802.15.4 slotted CSMA/CA MAC for time-critical events in wireless sensor networks, in: 5th international workshop on Real-Time Networks (RTIno6), Dresden, Germany, 2006.


M. Aykut Yigitel is currently pursuing his Ph.D. degree in Department of Computer Engineering, Bogazici University, Istanbul, Turkey. He received his B.S. and M.S. degrees in Computer Engineering from Hacettepe University Ankara, Turkey and Bogazici University, Istanbul, Turkey, in 2005 and 2010, respectively. He is also working at Turkish War Colleges Command as Wargaming System Manager. His research interests include design and performance evaluation of communication protocols for wireless ad hoc and sensor networks, and QoS provisioning in sensor networks.
Özlem Durmaz Incel is currently a post-doctoral researcher in the Networking Laboratory (NETLAB) of the Bogazici University, Turkey. She received her Ph.D. in computer science from the University of Twente, Netherlands, in March 2009. Her dissertation focused on efficient data collection in wireless sensor networks and was entitled as “Multi-Channel Wireless Sensor Networks: Protocols, Design and Evaluation”. She was a visiting student in the Autonomous Networks Research Group of the University of Southern California as part of her Phd studies in 2007–2008. She received both her MSc and BSc degrees in computer engineering from the Yeditepe University, Turkey, in 2005 and 2002, respectively. Her research interests are in the design and analysis of algorithms/protocols for wireless networks, particularly for ad hoc and sensor networks, and in the performance evaluation of computer networks.

Cem Ersoy received his B.S. and M.S. degrees in Electrical Engineering from Boğaziçi University in 1984 and 1986, respectively. He worked as an R&D Engineer at NETAS A.S. between 1984 and 1986. He received his Ph.D. in Electrical Engineering from Polytechnic University, Brooklyn, New York, in 1992. Currently, he is a professor in the Computer Engineering Department of Boğaziçi University. His research interests include performance evaluation of communication networks, wireless sensor networks, and mobile applications. He is the chairman of the IEEE Communications Society Turkish Chapter.